Geophys. J. Int. (2007) **169,** 951–963 doi: 10.1111/j.1365-246X.2007.03354.x

An improved geomagnetic data selection algorithm for global geomagnetic field modelling

Alan W. P. Thomson[1] and Vincent Lesur[1,2]

[1]*British Geological Survey, West Mains Road, Edinburgh* EH9 3LA, *UK. E-mail: awpt@bgs.ac.uk*
[2]*GeoForchungsZentrum, Potsdam, Germany*

Accepted 2007 January 10. Received 2007 January 10; in original form 2006 September 21

SUMMARY

A spherical harmonic degrees 60, global internal field model is described (called BGS/G/L/0706). This model includes a degree 15 core and piecewise-linear secular variation model and is derived from quiet-time Ørsted and CHAMP satellite data sampled between 2001.0 and 2005.0. For the satellite data selection, a wide range of geomagnetic index and other data selection filters have been used to best isolate suitably quiet magnetospheric and ionospheric conditions. Only a relatively simple, degree one spherical harmonic, external field model is then required. It is found that a new 'Vector Magnetic Disturbance' index (*VMD*), the existing longitude sector *A* indices, the auroral zone index *IE*, and the polar cap index *PC* are better than *Kp* and *Dst* at rejecting rapidly varying external field signals at low, middle, auroral and polar latitudes. The model quality is further enhanced by filling spatial and temporal gaps in the quiet data selection with a second selection containing slightly more disturbed data. It is shown that *VMD* provides a better parametrization than *Dst* of the large-scale, rapidly changing, external field. The lithospheric field model between degrees 16 and 50 is robust and displays good coherence with other recently published models for this epoch. BGS/G/L/0706 also shows crustal anomalies consistent with other studies, although agreement is poorer in the southern polar cap. Intermodel coherency reduces above about degree 40, most likely due to incompletely filtered signals from polar ionospheric currents and auroral field aligned currents. The absence of the *PC* index for the southern hemisphere for 2003 onwards is a particular concern.

Key words: crustal structure, geomagnetic field, inverse theory, lithosphere.

1 INTRODUCTION

The Ørsted and CHAMP magnetic survey satellites were launched in 1999 and 2000, respectively, and have since provided a largely uninterrupted set of vector and total intensity magnetic field data for global field modelling purposes. Even with such a data set there is a major problem to overcome in producing robust models of the internal geomagnetic field with fine spatial and temporal details. This is the difficulty in accounting for short wavelength and rapidly time-varying external fields, particularly those fields generated in the ionosphere and in the inner magnetosphere at high latitudes.

To circumvent this problem one approach is to process (sometimes referred to as filtering) the survey data, before the modelling stage, to remove, as far as is possible, contributions from magnetic sources not included in the model. For this processing method, models of unwanted source fields may be employed. The MF4 model (Maus *et al.* 2006a) is a well known example of such an approach. MF4 represents the field generated in the lithosphere and includes all spherical harmonic (SH) degrees from 16 to 90. However, there is a risk associated with data processing and this is that part of the modelled lithospheric field may be removed from the data, at the same time as the signals from the other, unmodelled, sources. Another recent model where data processing is involved is POMME-3 (Maus *et al.* 2006b).

Another, probably more common, approach seeks to minimize the contributions of unmodelled fields by careful data selection, using geomagnetic index and other activity data. Models are then constructed for all the signals believed to remain in the data, or an appropriate statistical treatment is used for unmodelled sources. The 'Comprehensive Model' (Sabaka *et al.* 2004), of degree 60, is one example of this approach. Another example is CHAOS (Olsen *et al.* 2006). The CHAOS model extends to SH degree 50 and contains a relatively complex magnetospheric field parametrization. The CHAOS model's other main characteristic is that the Euler alignment angles, relating the vector magnetometer and the star imager, are co-estimated with the field model. (We should note that both MF4 and POMME-3 also include a rigorous data selection, as part of their processing/filtering procedures.)

Up until now the 'data selection' (i.e. second) approach has probably not been as successful as the more complicated 'data processing'

(first) approach in modelling the lithospheric part of field, certainly at SH degrees above about 30–50. The main reason for this is the complexity, spatially and temporally, of the magnetic field generated by the system of currents in the polar ionosphere and magnetosphere. We simply lack sufficiently sophisticated ionospheric and magnetospheric field models and these signals, therefore, leak into combined core and crustal field models.

Most recent magnetic field models produced from the Ørsted and CHAMP data sets include a parametrization of the field generated by the magnetospheric ring current, usually based on the *Dst* index. *Dst* was first used by Langel and coworkers (e.g. Langel & Estes 1985, also see Sugiura 1964, regarding *Dst*), in constructing global field models from Magsat satellite data. *Dst* is derived from four low-latitude observatory data time-series, at a 1 hr time resolution, and the data are carefully processed to provide a quantitative measure of magnetic disturbance. A recent attempt to improve the parametrization of the ring current was provided by Olsen (2002), who used data from 50+ Intermagnet observatories to construct the '*RC*' index. [We should also note the potential usefulness of the *SYM-H* and *ASYM-H* indices as ring current monitors, for example, Iyemori (1990), although we do not consider these in this paper.] The assumption behind both the *Dst* and *RC* modelling approaches is that rapid variations of the inner magnetospheric field are mainly due to a symmetric ring current, where the perturbations are aligned with the internal dipole axis. However, it is now recognized that fields generated by other magnetospheric sources can be discerned in both the satellite and observatory data (Lesur *et al.* 2005a) and the contributions from these other fields have begun to be included in models (Lesur *et al.* 2005b; Maus & Lühr 2005). The modern use of long time-series of satellite data has also revealed other weaknesses of *Dst*, for example, uncertainties in its baseline stability and in the relative contribution from internal (induced) fields (Olsen *et al.* 2005; Maus & Weidelt 2004). To attempt to overcome these difficulties, we therefore, introduce a new 'Vector Magnetic Disturbance' (*VMD*) time-series, at a 20 min resolution, computed from low and mid-latitude observatory data.

Dst, as well as *Kp*, is often also used in the 'data selection' approach, where one selects magnetic survey data only during local night time, and where the contemporary *Kp* and *Dst* indices are small. For the most part these limited criteria have proved appropriate for models based on data of the quality of, for example, Magsat (e.g. Langel & Hinze 1998). However, in the Ørsted and CHAMP satellite era the data quality has proved to be considerably better and perhaps other or new indices may be more appropriate.

Certainly, in the present era, there are many more, readily available, geomagnetic and solar wind activity indicators. Many of these data are local in the sense that they may be closer to the position and time of the high-precision satellite magnetic measurements, certainly in comparison with global indices such as *Kp* and *Dst*. For example, there are the (3-hr) longitude-sector *A* indices, (Menvielle & Berthelier 1991; Menvielle & Berthelier 1992), the (5- and 15-min) polar cap *PC* index (Troshichev *et al.* 1988), and the (1-min) IMAGE auroral zone magnetometer array index *IE* (Viljanen & Hakkinen 1997) (the latter substituting for the currently unavailable northern hemisphere *AE* index). There are also estimators of auroral zone field aligned current strength (*FAC*), such as the time-rate-of-change of the dipole-East component of the field (Stauning *et al.* 2003), and the new (*VMD*) time-series.

To add to the wealth of information on local and global magnetic activity levels, we also have access to near-Earth solar wind data, in the form of the hourly interval OMNI data set (King & Papitashvili 2004) and the 1-min fully calibrated data from the NASA/NOAA ACE spacecraft (Garrard *et al.* 1998). These data, suitably delayed for propagation from the measurement point to the inner magnetosphere, provide information on energy input to the magnetosphere and on the solar wind electric field that drives the magnetospheric current circulation.

In this paper, we present an internal magnetic field model to SH degree 60 (known as BGS/G/L/0706, where *BGS* denotes 'British Geological Survey', *G* denotes 'global' and *L* denotes 'lithosphere' – to indicate the main purpose of the model – and dated 2006 July). To achieve this relatively high degree we have carried out an exhaustive re-examination of the value of both new and old data selection algorithms for satellite data. We have also adopted a new parametrization of the large scale symmetric external field, for example, replacing *Dst* by *VMD*. The *VMD* index is described in the next section. Section 3 then describes the data selection process. In Section 4, we discuss the various model parametrizations employed and in Section 5 we discuss some results arising from these models. We also present detailed results for the final model BGS/G/L/0706. Section 6 then concludes with some observations on possible future directions for this work.

2 THE 'VECTOR MAGNETIC DISTURBANCE' (*VMD*)

INTERMAGNET (http://www.intermagnet.org) 1-min averaged geomagnetic observatory data are processed to construct a time-series, the 'Vector Magnetic Disturbance' (*VMD*), that can be used to monitor rapid variations in both the strength and direction of the large-scale external field. *VMD* is produced as two 3-D vector time-series covering 3 month intervals: one 3-D time-series for the external field and one 3-D time-series for the associated induced field. Each time-series covers (for a given year) the months January–March, April–June, July–September, and October–December.

Data from all INTERMAGNET observatories between $+50^\circ$ and -50° magnetic latitudes are selected in the shadow zone of the Earth, that is, for local times (LT) from 23:00 to 05:00 and for zenith angles (z_a) larger than 101.54° [i.e. $\cos(z_a) < -0.2$]. Under these selection criteria, it is assumed that over a period of 3 months the 3-D data vector $\mathbf{d}_i(t)$ of observatory i is well described, in a geocentric system of coordinates, by:

$$\mathbf{d}_i(t) = c_i + t v_i + \mathbf{r}_i(t) + \mathbf{n}_i, \tag{1}$$

where $\mathbf{c}_i$, $\mathbf{v}_i$, $\mathbf{r}_i(t)$ and $\mathbf{n}_i$ are all 3-D vectors. $\mathbf{c}_i$ is a constant to account for the average internal and external field at the observatory site. The linear variation in time, t, of size $\mathbf{v}_i$, accounts for the slow time variations of the core field over the 3 month period. This linear variation in time may also describe a linear trend in the average external field. The vector $\mathbf{r}_i(t)$ describes the rapid time variations that we will assume to be due exclusively to the large scale (i.e. spherical harmonic degree 1) external field and its induced internal counterpart. Any signal that does not fit in this description is regarded as part of the noise, $\mathbf{n}_i$.

Estimating the *VMD* time-series is a three step process where first we estimate the observatory offsets $\mathbf{c}_i$ and linear trends $t\mathbf{v}_i$ [for this a relatively smooth model is required for $\mathbf{r}_i(t)$]. For the second step, with offsets and linear trends known, an accurate model of the time behaviour of $\mathbf{r}_i(t)$ is built but, due to the observatory distribution, the large scale external field signal cannot be well separated from its associated internal induced signal. This is done at the final step. The three steps are described in some detail below. Throughout this Section we use a Cartesian coordinate system with origin at the Earth's centre, Z parallel to the axis of rotation, X pointing

towards the Greenwich meridian, and Y completing a right-handed orthogonal coordinate system.

2.1 Step 1: Estimating the offset and linear trend

In step 1, we select data from magnetically calm days in 3 month sections (by rejecting 1 min means when *Dst* lies outside the range $[-20,\ 20]$ nT). These data are then fitted with the model given in eq. (1). $\mathbf{r}_i(t)$ is then the negative gradient of a potential: $\mathbf{r}_i(t) = -\nabla\rho(\theta, \phi, r,\ t)$, where

$$\rho(\theta,\phi,r,t) = a\sum_{m=-1,0,1}\left[a_m(t)\left(\frac{a}{r}\right)^2 + b_m(t)\left(\frac{r}{a}\right)\right]Y_1^m(\theta,\phi) + a\sum_{m=-1,1}\left[\alpha_m(t)\left(\frac{a}{r}\right)^2 + \beta_m(t)\left(\frac{r}{a}\right)\right]Y_1^m(\theta,\tilde{\phi}) \quad (2)$$

and where θ and ϕ are the colatitude and longitude in geocentric system of coordinates and $a = 6371.2$ km is an average Earth radius. The $Y^m{}_1(\theta, \phi)$ are the degree 1 Schmidt normalized spherical harmonics with the convention that the $m \geq 0$ terms are associated with cos $(m\,\phi)$ and the $m < 0$ terms are associated with sin $(m\,\phi)$. The angle $\tilde{\phi}$ is the local time of the observatory in degrees (i.e. $\tilde{\phi} = \phi + \frac{360}{24}t$, t being the universal time in decimal hours, and ϕ the longitude of the observatory). The time dependence of the $a_m(t)$, $b_m(t)$, $\alpha_m(t)$ and $\beta_m(t)$ is defined by B-splines with a node every 3 days. This node separation has been chosen as a trade-off between a better fit to the data and a robust estimation of the $\mathbf{c}_i$ and $\mathbf{v}_i$. The potential $\rho\ (\theta, \phi, r, t)$ has a very smooth behaviour in time but, owing to its $\tilde{\phi}$ dependence, it is nonetheless able to fit magnetic fields that have an apparent 24 hr periodicity; for example, the periodicity due to the rotation of the Earth in a slowly varying, Sun-fixed external magnetic field (Lesur *et al.* 2005b; Maus & Lühr 2005).

2.2 Step 2: Estimating the combined internal and external *VMD* series

In step 2, we consider a second data set, similar to the above but without any *Dst* activity filtering. The offsets ($\mathbf{c}_i$) and linear trends ($t\mathbf{v}_i$) are subtracted from each observatory component. The time span of data is also extended backward by 1 month such that 4 months of data are available. We then fit this extended data set by a large-scale internal and external magnetic field with the associated magnetic potential:

$$\tilde{\rho}(\theta,\phi,r,t) = a\sum_{m=-1,0,1}\left[\tilde{a}_m(t)\left(\frac{a}{r}\right)^2 + \tilde{b}_m(t)\left(\frac{r}{a}\right)\right]Y_1^m(\theta,\phi) \quad (3)$$

The time dependence of $\tilde{a}_m(t)$ and $\tilde{b}_m(t)$ is now defined by B-splines with one node per hour. The distribution of observatories over the Earth is such that there is not a robust separation of internal and external contributions and it is, therefore, necessary to recombine these contributions and define the *VMD* as:

$$VMD_m(t) = -\tilde{a}_m(t) - \tilde{b}_m(t) \quad (4)$$

The $VMD_m(t)$ as defined by eq. (4) is the combined SH degree 1 external and internal magnetic field that would be measured on the Earth's equator in a geocentric Cartesian system of coordinates, where $m = -1,\ 0,\ 1$ correspond to the Y, Z and X directions, respectively.

2.3 Step 3: Separating internal and external contributions in *VMD*

The $VMD_m(t)$ defined in eq. (4) is then sampled every 20 min over each full 4 month period and de-trended. The separation of internal from external contributions is then done in the same manner as described in Olsen *et al.* (2005) and Maus & Weidelt (2004). The reference model we use for this purpose is that constructed by Utada *et al.* (2003). Olsen's filter Q_1 (Olsen *et al.* 2005), is computed in the time domain over a 30 days period such that, when convolved with the 4 month $VMD_m(t)$ time-series, it results in the desired 3 month interval of large-scale-induced internal perturbations: $VMD_m^{\rm int}(t)$. Then the external part $VMD_m^{\rm ext}(t) = VMD_m(t) - VMD_m^{\rm int}(t)$ is derived for the same 3 month interval. Fig. 1 shows an example of $VMD_m^{\rm ext}(t)$, in this case for July to September 2001.

3 DATA SELECTION

Global geomagnetic indices such as *Kp* and *Dst* are commonly used to isolate intervals of quiet-time magnetic data for internal field modelling, where the magnetospheric field is relatively simple, large scale (e.g. describable by low degree SH) and quasi-static. However, there are many other indices of activity available, both geomagnetic and solar wind, and either global or local in applicability.

For this paper, we have carried out a re-appraisal of many of these indices and solar wind data. We have then used the most appropriate combinations of these indices to isolate the quietest satellite magnetic data in the Ørsted and CHAMP data sets, between 2001.0 and 2005.0. In Table 1, we list all the indices that have been studied, together with some comments, and present the final (BGS/G/L/0706) data selection filters, as chosen by experimentation. The most appropriate indices and their pass bands were determined from an examination of the spectral power (Lowes 1966), using the model parametrizations of Section 4, with the simple aim of minimal spectral power in the internal field.

We are unsurprisingly unable to establish 'best' filter combinations. The BGS/G/L/0706 filter cut-offs and index types are trade-offs between the different combinations of parameters available (realistically, the parameter space is too large to study in a systematic way). It may therefore, be that for different data sets, or for a different modelling purpose, there may be other, more appropriate, filters and passbands and that these should be found by experiment.

The BGS/G/L/0706 data sampling was one scalar ($>|50|^\circ$ magnetic latitude) or vector sample ($\leq|50|^\circ$ magnetic latitude) per 20 s, selected according to the filters of Table 1, as indicated by a '♯'. Data for other models (i.e. those constructed for relative comparisons and discussed in Section 5) were derived in a similar manner. Each scalar sample was replaced by a 'projected-F' value, that is by the component along an estimated field direction at the sample point (the dot product of the measured vector field and an *a priori* unit field vector). We used an existing simpler BGS global field model for the same period to estimate this *a priori* local field direction. This approach is used because it makes the inverse problem linear.

The IMAGE magnetometer array index *IE* Viljanen & Hakkinen (1997) only covers a restricted window of local time ($\sim$2 hr), in comparison with the hemispheric *AE* index. For that reason we also studied local estimates of the field aligned current (*FAC*) density at the geographic location of the satellite sample, as a possible additional filter, following Stauning *et al.* (2003). Stauning *et al.*'s prescription provides an approximate current strength for polar orbiting satellites from the geomagnetic east component of the field: $FAC \cong 0.1\frac{\partial By}{\partial t}$ (μA m^{-2}). As a selection parameter we find that

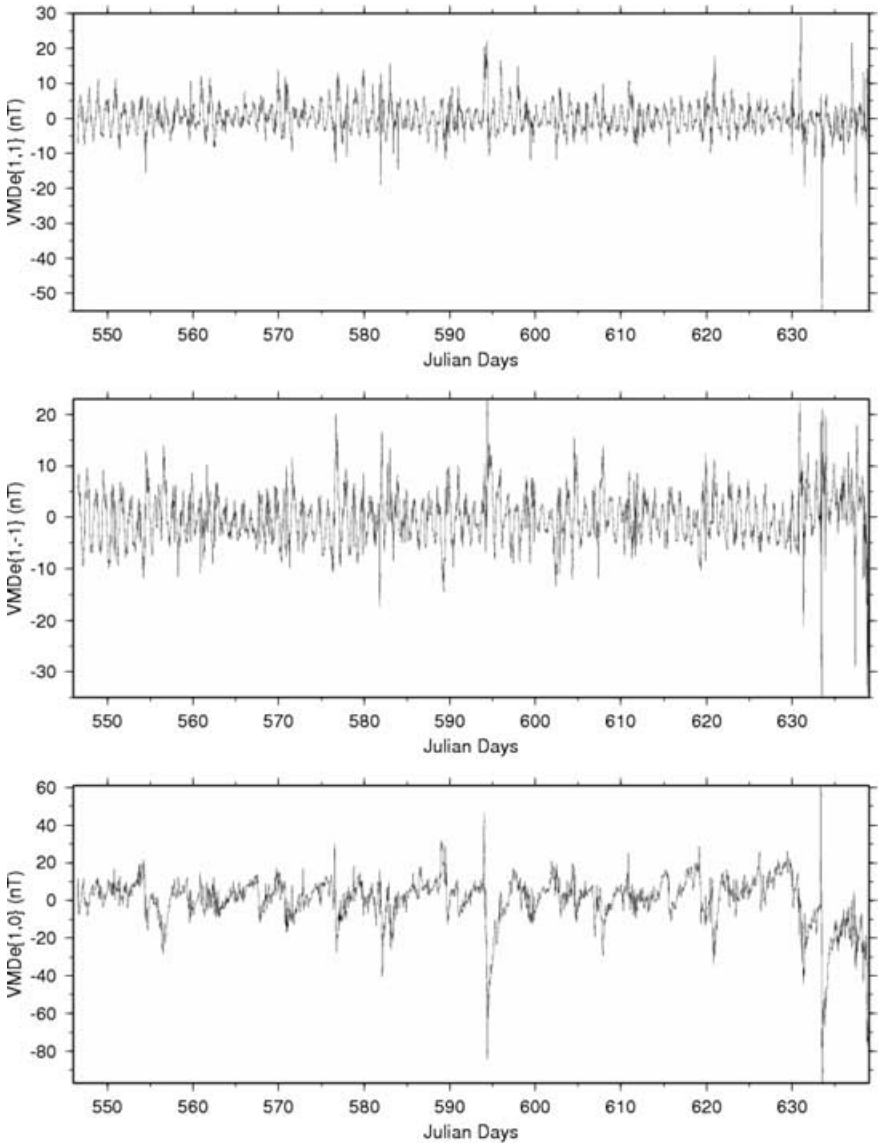

Figure 1. The three components of the external *VMD* index for 2001 July–September, as a function of Julian day. The (1, 0) term, aligned with the rotation axis, closely resembles the hourly *Dst* index, which is more aligned with dipole north, but it has more structure due to its 20-min sample rate. The other terms (*top* = *X*, *middle* = *Y*, in a Cartesian system) show a strong 24 hr periodicity.

choosing quiet levels of *FAC*, say $\leq 0.8\ \mu\mathrm{A\ m^{-2}}$ (Ritter *et al.* 2004), leads to probably too few auroral zone data. We feel that further study of this parameter is needed and we, therefore, do not use the *FAC* proxy in producing BGS/G/L/0706. Some further remarks on this proxy are made in Section 6.

The PC_S index for the southern polar cap is not currently available for 2003 and onwards. In this case, we have simply used the Northern hemisphere PC_N to select post-2003 southern polar data. This is clearly not ideal and we show later that it does cause problems with field models for Antarctica and highlights the need for an update of this valuable time-series.

We note that although we tested the solar wind merging electric field parameter, E_m, of Ritter *et al.* (2004) we probably used too wide a filter (a value suggested by the results of Ritter *et al.* 2004) and we, therefore, do not prefer it here to *PC*. However, as with the *FAC* density proxy, further study is probably warranted.

The filters and pass bands were determined from relative comparisons between models constructed on a shorter span of data, 2001.0-2002.5, simply for processing convenience. Whilst this period provides an excellent temporal and spatial density of quality data, the optimum index filters (i.e. those '♯' in Table 1) then produce a sparser data set when applied during the other years (i.e. post 2002.5) for BGS/G/L/0706. For this reason we have supplemented these 2001.0–2005.0 data with a second, more weakly filtered, data set. The data in this second set are generally more active, as indicated in the final column of Table 1. These secondary data were selected only in those rectangular tesserae (equal to 1° square at the equator) where there were no data present from the first round of selection. We have called this process 'iterative data selection'.

Fig. 2 shows the numbers of data in each tesseral bin for, respectively, the 1st and 1st + 2nd rounds of iterative data selection, for 2001.0–2002.5. The selection criteria are given by the '♯' filters of Table 1. We note that iterative data selection is applied separately for each year over 2001–2005 for BGS/G/L/0706, to maximise the evenness of the data distribution in time. We also note that it may be possible to determine an appropriate weight for the second data set

Table 1. Geomagnetic index, solar wind data and other filters used in data selection. A ♯ denotes those data types used to select data for the final model BGS/G/L/0706, using first the 'strong filter', with the 'weak filter' subsequently used for data gap filling. IMF is the Interplanetary Magnetic Field, in Geocentric-Solar-Magnetospheric (GSM) coordinates; solar wind speed is V_{SW}, and the 'pre-model residual' is with respect to a prior *BGS* global internal model, of degree 20, based on Ørsted and CHAMP data for a similar time span. Note that we also require equally quiet *Kp* and Sector-*A* data for the previous 3 hr interval, and equally quiet *VMD* (or *Dst*) during the previous 1 hr. The 'weak filter' is discussed below, in relation to the iterative data selection process.

Data type	♯ = Used in final selection	Strong filter passband	Weak filter passband
D_{st}		$[-20, +20]\ nT$	$[-20, +20]\ nT$
$\lvert\frac{\partial D_{st}}{\partial t}\rvert$	♯	7 nT h^{-1}	20 nT h^{-1}
VMD		$[-20, +20]\ nT$	$[-20, +20]\ nT$
$\lvert\frac{\partial VMD}{\partial t}\rvert$		7 nT h^{-1}	20 nT h^{-1}
Kp		$\leq 2-$	$\leq 2+$
Sector-*A*	♯	$\leq 2-$	$\leq 2+$
IE	♯	[0, +30] nT	[0, +100] nT
PC_N	♯	[0.0, +0.2] mV m^{-1}	[0.0, +1.0] mV m^{-1}
PC_S	♯	[0.0, +0.2] mV m^{-1}	[0.0, +1.0] mV m^{-1}
Em		[0.0, +0.8] mV m^{-1}	[0.0, +2.0] mV m^{-1}
FAC		$[-0.8, +0.8]\ \mu A/m^2$	$[-1.5, +1.5]\ \mu A/m^2$
Solar Zenith	♯	Dark up to 250 km (applied only in polar cap)	Dark up to 250 km
Local Time	♯	22:00−05:00 (not used in polar cap)	22:00−05:00
IMF B_Z	♯	[0, +6] nT	[0, +10] nT
IMF B_Y	♯	[−3, +3] nT	[−10, +10] nT
IMF B_X		[−10, +10] nT	[−10, +10] nT
V_{SW}	♯	[0, 450] km s^{-1}	[0, 450] km s^{-1}
Pre-Model |*Residual*|	♯	2sigma max. =10 nT	4sigma max. =20 nT

(rather than the unit weight assumed here), through a trade-off of model quality versus the weight (N. Olsen, personal communication, 2006).

4 MODEL PARAMETRIZATION

The magnetic field model, away from sources, is the negative gradient of a potential: $\mathbf{B} = -\nabla\mathbf{V}(\theta, \phi, \mathbf{r}, t)$. The potential dominated by contributions from sources in the liquid outer core and the permanently magnetised crust of the Earth is defined by:

$$V^{\text{int}}(\theta,\phi,r,t) = a\sum_{l=1}^{15}\sum_{m=-l}^{+l} g_l^m(t)\left(\frac{a}{r}\right)^{l+1} Y_l^m(\theta,\phi) + a\sum_{l=16}^{60}\sum_{m=-l}^{+l} g_l^m\left(\frac{a}{r}\right)^{l+1} Y_l^m(\theta,\phi), \quad (5)$$

where a and $Y_l^m(\theta,\phi)$ are defined in Section 2. The magnetic potential for large-scale external sources is defined as:

$$V^{\text{ext}}(\theta,\phi,r,t) = a\sum_{m=-1}^{m=+1} q_1^m(t)\left(\frac{r}{a}\right) Y_1^m(\theta,\phi) \quad (6)$$

The time dependence of $g_l^m(t)$ and $q_1^m(t)$ is piecewise linear with nodes at 2001.5, 2002.5, 2003.5 and 2004.5. For data outside the range (2001.5, 2004.5) the model is extrapolated following the linear trend of the closest node interval. A more complex time variation, for example, cubic B-spline, may be preferable for secular variation modelling but the time span here is relatively short and the main emphasis is on the static, high degree, field. The above potentials do not include any *VMD* dependence. We introduce *VMD* via the potential:

$$V^{\text{vmd}}(\theta,\phi,r,t) = a\sum_{m=-1}^{m=+1}\left\{\left[a_{1,m} + ta_{2,m} + a_{3,m}VMD_m^{\text{int}}(t)\right]\left(\frac{a}{r}\right)^2 + \left[b_{1,m} + tb_{2,m} + b_{3,m}VMD_m^{\text{ext}}(t)\right]\left(\frac{r}{a}\right)\right\} Y_1^m(\theta,\phi) \quad (7)$$

Between 2001.0 and 2005.0, 16 different $V^{\text{vmd}}(\theta,\phi,r,t)$ potentials are defined, one for each successive 3 month period, and without any continuity constraints imposed between them.

To have a good representation of the 24 hr, semi-annual and annual periodicity of the external field, we also introduce a potential defined as:

$$V^{\text{ann}}(\theta,\phi,r,t) = a\sum_{m=-1}^{m=+1}\left[i_m(t)\left(\frac{a}{r}\right)^2 + e_m(t)\left(\frac{r}{a}\right)\right] Y_1^m(\theta,\phi) + a\sum_{m=-1}^{m=+1}\epsilon_m(t)\left(\frac{r}{a}\right) Y_1^m(\theta,\tilde{\phi}) \quad (8)$$

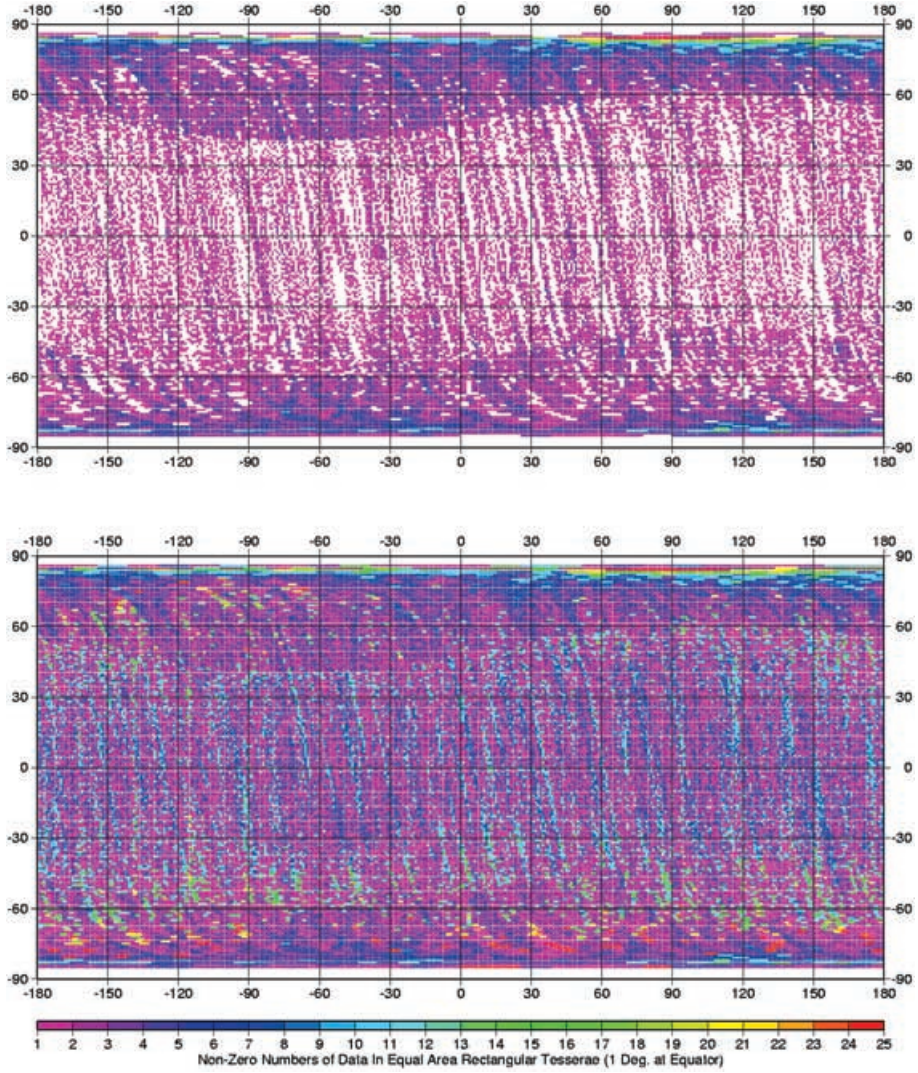

Figure 2. Numbers of data per 1° latitude–longitude tessera in the 1st (top: 'strong filter') and second round (bottom: 'weak filter'; with data added to the 1st round of data selection) of the iterative data selection process. Ørsted data for 2001.0–2002.5 only are shown in this example. White space indicates no data.

where $\tilde{\phi}$ is defined in Section 2. The time dependency for $i_m(t)$ is: $i_m(t) = i_m^c \cos(2\pi t) + i_m^{2c} \cos(4\pi t) + i_m^s \sin(2\pi t) + i_m^{2s} \sin(4\pi t)$ where the time is given in decimal year. An equivalent time dependence holds for $e_m(t)$. The time dependency for $\epsilon_m(t)$ is: $\epsilon_m(t) = \epsilon^0 + \epsilon_m^c \cos(2\pi t) + \epsilon_m^s \sin(2\pi t)$. [We did not introduce an internal counterpart corresponding to $\epsilon_m(t)$ as we were not able to separate it properly from the other model parameters.] There is obviously some redundancy in this parametrization as the average of the $a_{1,m}$ and $a_{2,m}$ cannot be separated from the $g_1^m(t)$ in $V^{\text{int}}(\theta, \phi, r, t)$ and similarly for the external part.

Over the 4-yr period, the number of parameters for the internal contribution (excluding induced fields) is 4485 whereas for external contributions (including induced fields) the number is only 330 of which 288 are associated with the internal and external *VMD* time-series.

In Section 5 results are presented that use a *Dst* parametrization of the magnetospheric field (to support the case for *VMD*). In this case, $VMD_m^{\text{ext}}(t)$ and $VMD_m^{\text{int}}(t)$ in eq. (7) are replaced by $Dst^{\text{ext}}(t)$ and $Dst^{\text{int}}(t)$, as defined in Olsen *et al.* (2005).

5 RESULTS AND DISCUSSION

The appropriate set of data selection filters and the model parametrization for BGS/G/L/0706 were chosen by relative comparisons between different internal field models. CHAMP vector data were corrected for alignment error using the parameters provided by S. Maus (personal communications, 2005, 2006). Each of these models was derived from competing data selection options and parametrizations, with the aim being minimum spectral noise (as in Section 3). Our inverse modelling methodology involved an initial inversion using an L2 norm, followed by 25 iterations to implement an L1 norm (with no advantage seen to accrue from additional iterations). No damping or regularization was involved, though the data were (tesseral) weighted by latitude according to the method of Lesur *et al.* (2005a).

In Fig. 3, we show (top plot) one example of the relative impact in a short 2001.0–2002.5 data set, of different external model parametrizations and data selection options. Fig. 3 demonstrates the deterioration in internal field model quality as the external field

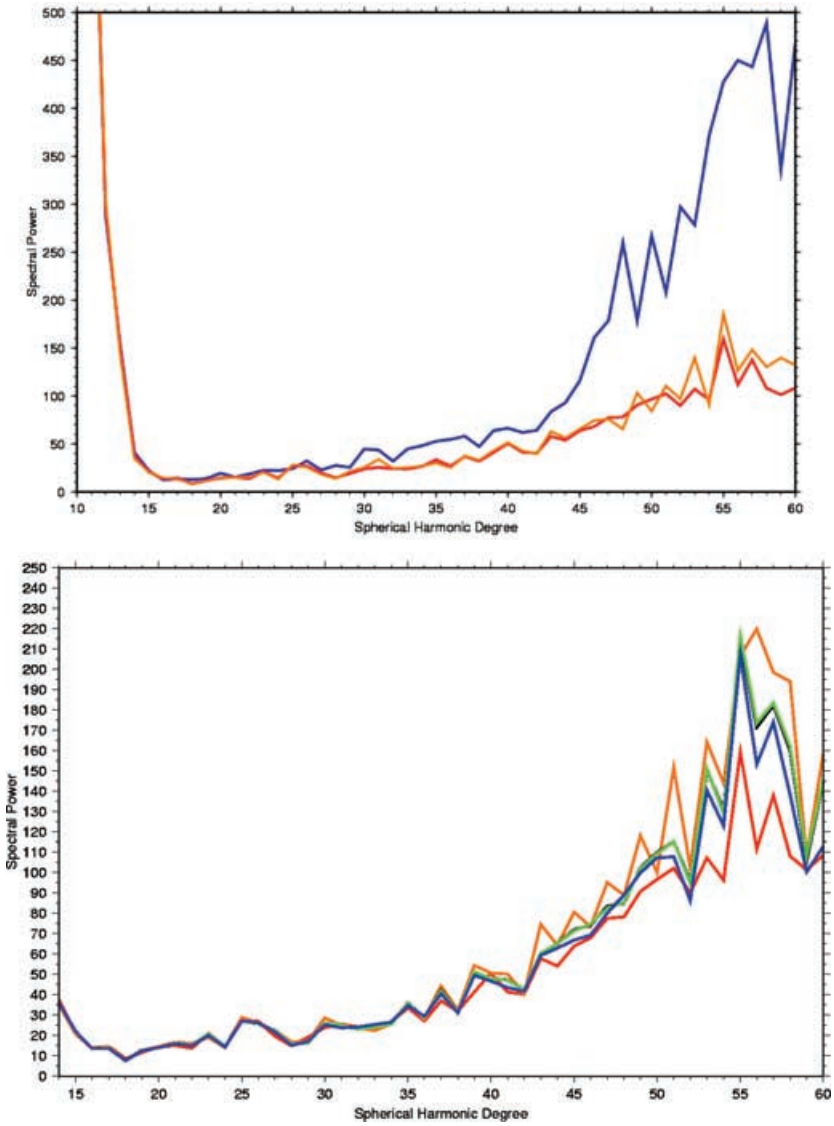

Figure 3. Top. Comparison of internal model power spectra for different external field parametrizations and with/without iterative data selection. The vertical scale is nT^2. The red line is the 'baseline' model, (Sections 3 and 4). Blue is a model derived from a similar data set but without iterative data selection (*cf*. upper panel only of Fig. 2). Orange is the spectrum from a model which has neither *VMD* or *Dst* dependence for the external field, but does have iterative data selection. Bottom. Comparison of internal model power spectra, for examples, of different data filters. Red denotes the baseline model (as top plot). Black, orange, green and blue then show the power spectra of models where the data selection has used, respectively, the Olsen-Maus decomposition of *Dst* (rather than *VMD*), the merging solar wind electric field (in place of *PC*), $dVMD/dt$ (in place of $dDst/dt$), and longitude sector *A*-indices (in place of *Kp*). Note: the 'baseline' model is derived from 20 per cent more data, hence the lower average noise level per degree, for example, above SH = 40.

component is simplified (in this case where there is no *VMD* or *Dst* dependence for the external SH=1 terms). The value of the iterative data selection is also clearly demonstrated by the significantly more noisy spectrum obtained from a data set with unfilled tesserae.

In Fig. 3, we also present some field model spectra to demonstrate the relative importance of different types of index filtering (lower plot). Different selection filters and pass bands usually give rise to different numbers of data, so intermodel comparisons can be difficult (e.g. more data should, in principle, produce a lower noise level). For example, the 'baseline' spectrum of Fig. 3 (lower plot), in red, is based on substantially more data than the other spectra and, unsurprisingly, shows a lower noise level particularly towards higher SH degree.

However, we can make sensible relative comparisons between the other models of Fig. 3 (lower plot). While these are also derived from different data sets (because we have used different index filters) they all have approximately the same numbers of data, to within a few per cent. Fig. 3 then suggests that there is probably no greater merit in filtering based on the merging electric field, compared to *PC*. It is possible that too high a filter cutoff has been used here and we have introduced more high latitude field noise. Use of $dVMD/dt$ is probably marginal compared to $dDst/dt$ as a data selection option (and we don't, therefore, use it in the final model), while there would seem to be merit, that is, a lower relative noise spectrum, in using the more localized activity monitor at mid-latitude, the longitude sector *A*-indices, compared to using global *Kp*.

VMD models typically result in smaller *rms* misfits compared to *Dst* based models (see Table 2). This does not lead to a significant improvement of the model itself as the data accumulation tends to average out the effect of the rapid variations of the large scale

Table 2. Residual average and *rms* misfit after the L2 norm inversion, for a model based on *VMD* or the *Dst* index. Field component residuals are given in spherical geocentric North/East/Down and in the SM system of coordinates. 'Project' denotes the scalar value in the direction of the prior model, that is, 'projected-*F*'.

Model type	Average (nT)				rms (nT)			
	North	East	Down	project.	North	East	Down	Project.
VMD	0.09	0.09	0.07	−1.54	4.39	4.41	3.45	9.79
	X (SM)	Y (SM)	Z (SM)		X (SM)	Y (SM)	Z (SM)	
	0.02	−0.07	0.01		3.38	4.11	4.73	
	North	East	Down	project.	North	East	Down	Project.
Dst	0.08	−0.01	0.03	−1.41	4.82	4.57	3.35	9.70
	X (SM)	Y (SM)	Z (SM)		X (SM)	Y (SM)	Z (SM)	
	0.03	−0.03	−0.01		3.39	4.22	5.10	

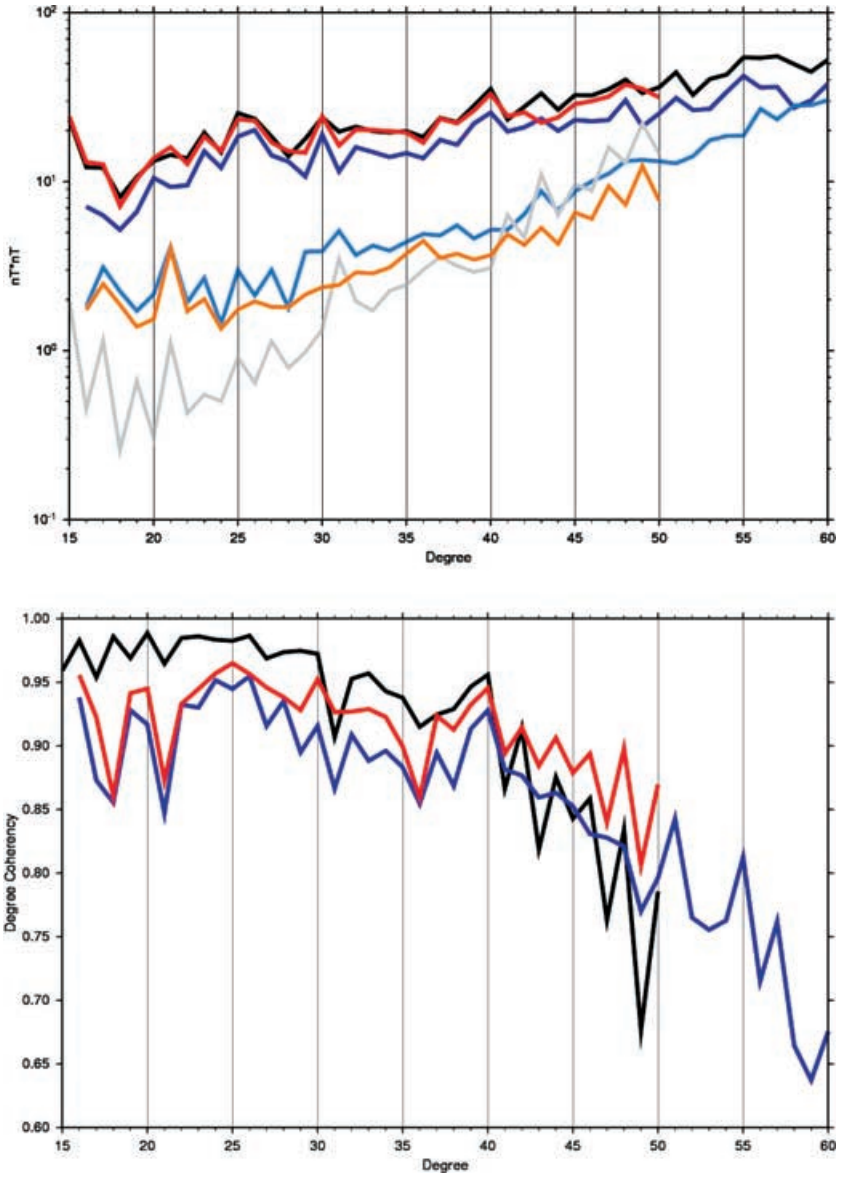

Figure 4. Top: The lithospheric power spectrum for BGS/G/L/0706 is shown in black. This is compared with CHAOS (red) and MF4 (blue). Power spectral differences per degree are then shown for BGS-CHAOS (grey), BGS-MF4 (pale blue) and CHAOS-MF4 (orange). The differences are constructed as the sum over all orders at each given degree, of the squared coefficient difference. Bottom: The degree coherency between the BGS model and CHAOS is shown in black, the coherency between BGS and MF4 in blue and that between CHAOS and MF4 in red.

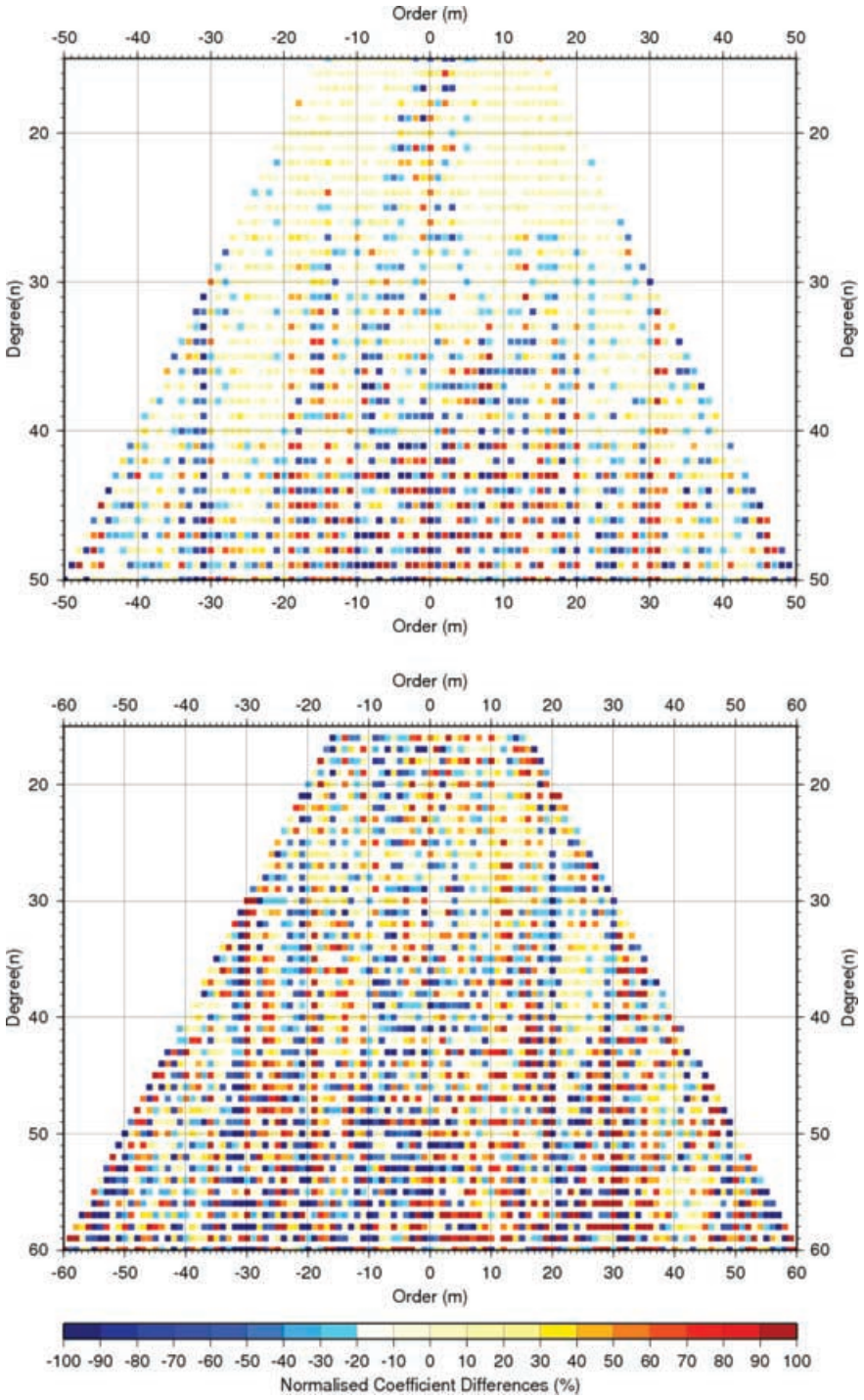

Figure 5. The normalized coefficient differences for BGS-CHAOS (top) and for BGS-MF4 (bottom).

external field. This latter field is, however, better described by using the *VMD* time-series than by using the *Dst* and that may have an impact for models built on short time scales. Improvements are clearly visible in the fit to the *Y* and *Z* components in the SM (solar-magnetospheric) coordinate system.

The BGS/G/L/0706 lithospheric spectrum (degrees 15–60) is shown in Fig. 4 (top), in comparison with spectra from both CHAOS (to degree 50) and MF4 (to degree 60). The BGS model, as with CHAOS, shows a distinct spectral dip at degree 18, as well as sharp peaks at degree 25, 30 and 40 (for example). Above degree 40 the BGS model is more similar to MF4 to around degree 55 (allowing for the obvious power level difference).

The degree coherency (Olsen *et al.* 2006), pairwise between the three models is shown in Fig. 4 (lower plot) and sensitivity matrices are provided in Fig. 5. We can see particularly good coherency between BGS/G/L/0706 and CHAOS up to spherical harmonic degree 30. There is then a fall-off in coherency from degrees 30–50, and BGS/G/L/0706 becomes more coherent with MF4 (without quite matching the coherence observed between CHAOS and MF4 for degrees 40–50). The BGS crustal model displays a slow 'roll-off'

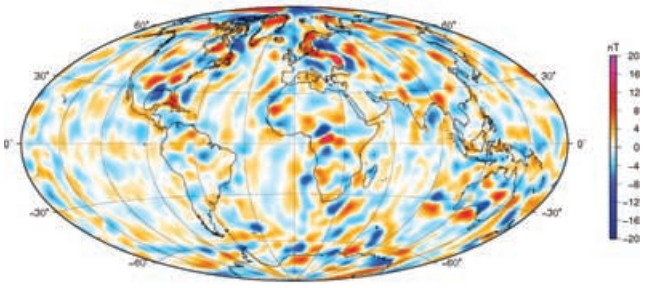

Figure 6. The lithospheric vertical field component (nT), from BGS/G/L/0706 at an altitude of 400 km, comprised of spherical harmonic degrees between 16 and 50.

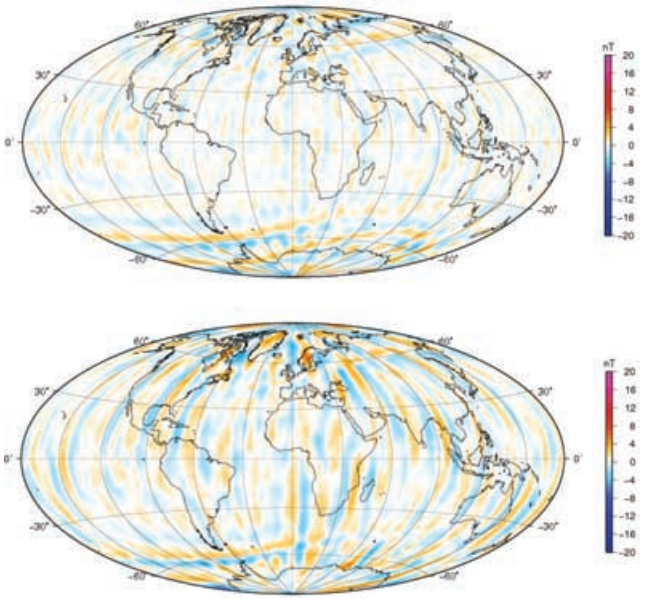

Figure 7. Vertical field component differences (nT) between BGS/G/L/0706 and CHAOS(top) and between BGS/G/L/0706 and MF4 (bottom), both at an altitude of 400 km and for spherical harmonic degrees between 16 and 50.

in coherency with MF4 (as opposed to the more rapid decay in coherency observed between BGS/G/L/0706 and CHAOS) from degree 40 to around degree 55. In terms of the sensitivity matrices (Olsen *et al.* 2006), strong zonal differences are probably the most obvious features (e.g. low order term differences between BGS/G/L/0706 and CHAOS below degree 30).

In Fig. 6, we show the BGS model lithospheric vertical field at a reference altitude of 400 km. In Fig. 7 the model differences BGS-CHAOS and BGS-MF4 are given. There are a number of significant features that can be observed. In Fig. 6 the major known crustal anomalies are clear, as well as the weaker oceanic signals. In Fig. 7, we see distinct meridional differences between the models. Probably the clearest feature however is the strong difference signal over the southern polar cap. This is not unexpected, being most likely due to the lack of PC_S data available to reject high southern latitude magnetic 'noise' for 2003 and onwards. We also use high latitude projected-F data in place of true total intensity data. Thus there are areas where the model needs further work, as we discuss further below.

6 CONCLUSIONS

BGS/G/L/0706 is the result of an extensive re-analysis of night-side Ørsted and CHAMP data, for 2001.0–2005.0. That we have been able to model the internal field to degree 60, and probably robustly to about degree 50, is mostly due to a careful search for appropriate satellite data selection techniques. In particular we have used novel magnetic indices not widely employed by the global field-modelling community. The external field model we have used is also relatively simple, for example, in comparison to CHAOS. We have also not utilized detailed data processing methodologies, for example, those used in the production of MF4.

We find that *VMD* improves the modelling of the external field, in comparison with *Dst*. *VMD* is intended to represent the large scale magnetospheric field but, in principle, it may contain any external signal, even those generated by *FAC*. BGS/G/L/0706 compares well with recent global internal models of high degree, showing a particularly good coherence with the CHAOS model to at least degree 30 and a moderately satisfactory coherence with MF4 to degree 60. Both the *VMD* index and the model coefficients can be obtained from the authors on request.

There are clearly areas that need further study, not least in the polar regions, where it is felt that the external field variations have probably been inadequately handled. Track-by-track analysis of polar passes to obtain quiet data is possible but it is undoubtedly time consuming, although it may prove necessary. However, this approach is at odds with the philosophy adopted in this paper, of automating the process as far as is practically possible. The southern polar cap is particularly problematic. Building high sample rate polar and auroral zone monitors, in a manner analogous to the *VMD* index, from high latitude INTERMAGNET observatory data, would seem to be another obvious way forward. The need for an updated PC_S, post-2003, is also clear from our results.

There are undoubtedly other improvements in data selection and modelling that are possible. Some preliminary ideas for future work, 'by-products' of the current research, are outlined here.

(i) Since the Ørsted and CHAMP data have been decimated by a factor of twenty, the variability ('noisiness' or 'uncertainty') of each data sample can be examined by computing a sample standard deviation (*SD*) over the 20-s decimation interval. In Fig. 8, we show the *maximum* sample standard deviation for the Y- (spherical geocentric east) component of the Ørsted measured field, binned in 1° tesserae, computed in magnetic local time (*MLT*) and dipole latitude. (The minimum sample standard deviation per bin is typically close to zero so we do not show it here.) The X- (north) component *SDs* are similar in magnitude, while the Z- (down) and F- (total field) component *SDs* are between 5 and 15 times smaller, respectively. In principle, the data of Fig. 8 are useful additional information for modelling, perhaps inverse-squared, as *a priori* data weights. Alternatively the *SD* data could be used as an additional selection principle, perhaps where the aim is attaining equal sample *SD*'s per tessera, across the globe.

(ii) As discussed earlier, the field-aligned current (*FAC*) density can be roughly estimated from the time-rate-of-change of the field component perpendicular to the main field and to the satellite ground track (Stauning *et al.* 2003). For satellites orbiting perpendicular to east–west, sheet-like, near vertical, *FAC* (high latitudes) we approximate the perpendicular component by the dipole East field component (Y). (The Y-component approximation is, therefore, less valid for low and mid-latitude data.) We show the average *FAC* (actually ≈10 *FAC*) in Fig. 9, derived for the Ørsted data set, where the rate-of-change is computed between consecutive 1-s samples, that

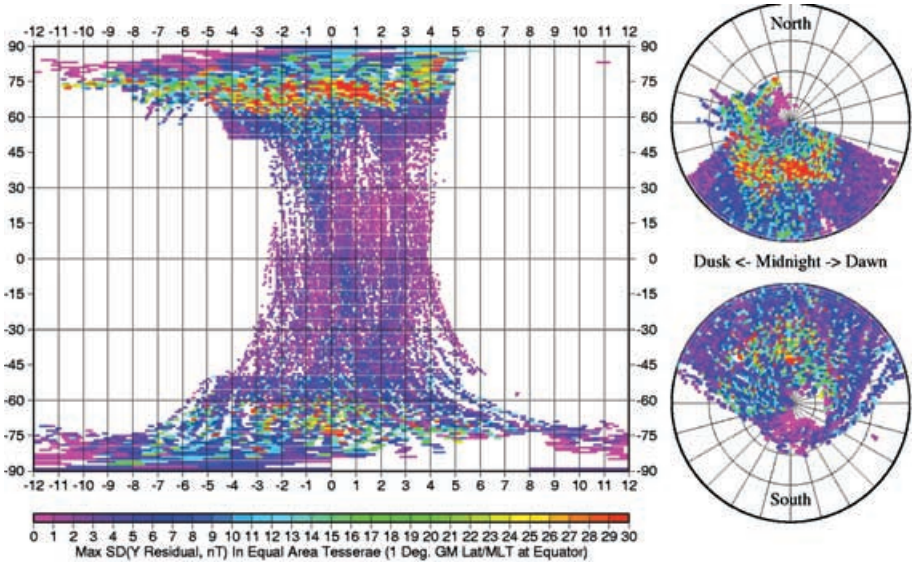

Figure 8. Maximum standard deviation in 1° tesseral bins for Ørsted Y-component data, used in the model, displayed in a magnetic local time (*MLT*) and dipole latitude system. Standard deviations are calculated from each consecutive sequence of 20 one-second samples, centred on the sample used in the modelling. The latitude-dependent distribution in *MLT* is a consequence of the local time and zenith angle filtering, as described in Section 3.

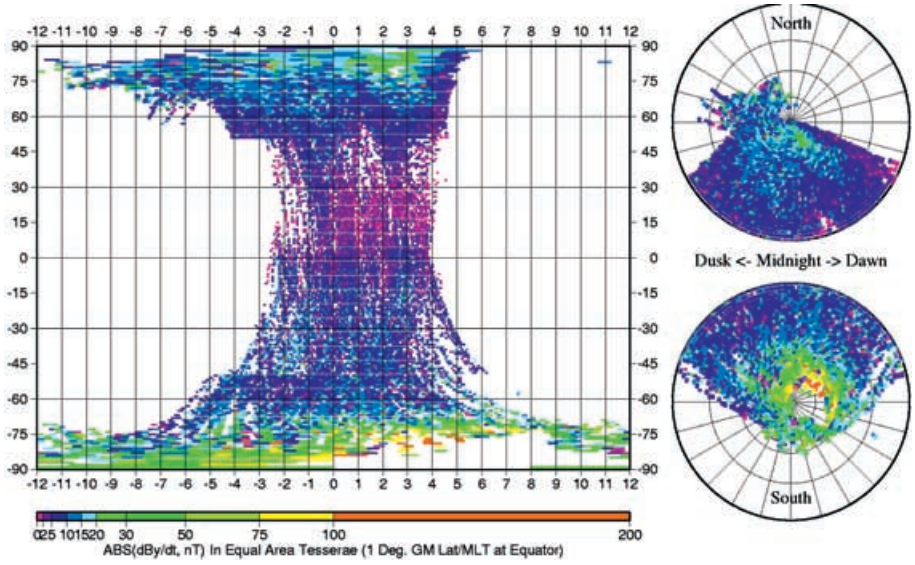

Figure 9. Average $|dBy/dt|$ in 1-deg tesseral bins for Ørsted data, as a proxy for the field-aligned current density, shown in *MLT* and dipole latitude coordinates.

is, with about 8 km of along-track separation. It is seen that the *FAC* distribution largely mirrors that of Fig. 8, but occurs towards higher latitudes (possibly Region 1 *FAC*). The Ørsted data also show a higher average *FAC* density compared to CHAMP, although the current morphology is otherwise rather similar. Whether this difference in density is significant we do not know, but we should note that CHAMP has lower and less anisotropic noise. As with Fig. 8, this *FAC* estimate could be used either in selection or in modelling. However, as we noted earlier, our initial attempts with this *FAC* measure proved unsatisfactory, at least for the purposes of BGS/G/L/0706.

(iii) The distribution of misfit (in the sense of data-(internal+external) model) for BGS/G/L/0706 is shown in Fig. 10. It is found that the different data selection filters (*cf.* Section 3) only marginally change this average misfit distribution, suggesting that unmodelled large scale external currents systems dominate. The large positive peaks in the polar regions could be the result of a pair of night-side solenoidal fields (i.e. the night-side tail+magnetopause current systems), with the field directed towards the Earth in the northern hemisphere and away from the Earth in the southern hemisphere. The magnitude of the misfit is certainly consistent with the estimated tail field at the Earth. The mid latitude, near midnight data, particularly in Z (i.e. positive towards the Earth in the north and negative in the south) also support these observations to some extent. However, more work is clearly needed on interpreting the misfit distribution and developing models of the tail and magnetopause current systems, for night-side data (Maus & Lühr 2005).

ACKNOWLEDGMENTS

We would like to acknowledge the Ørsted and CHAMP science data centres and the efforts of the many institutes that support the

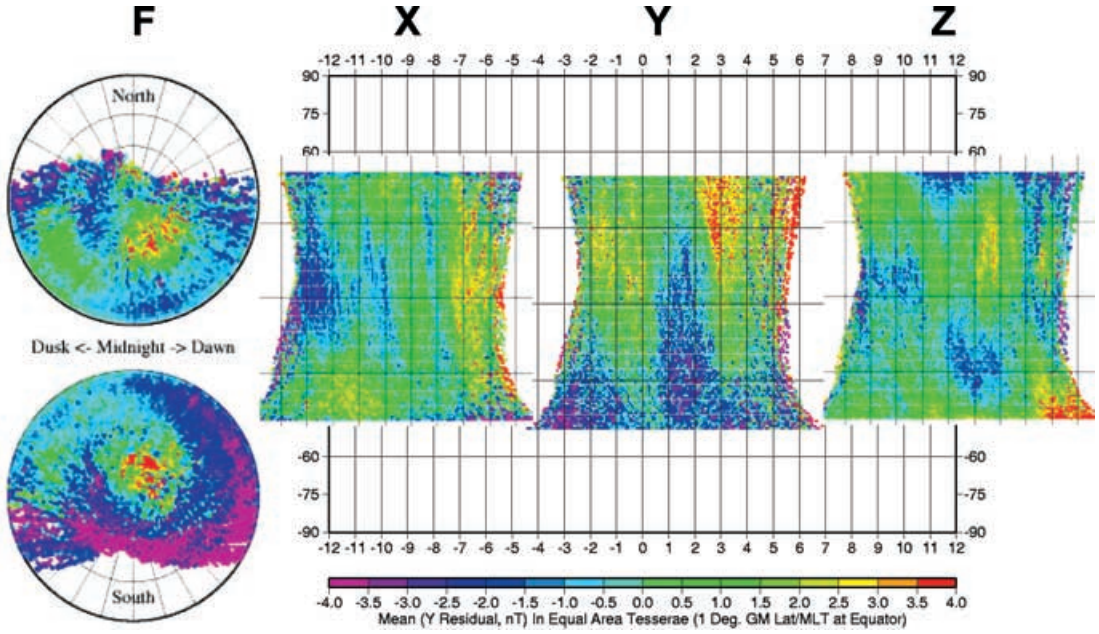

Figure 10. Average model misfit in 1° tesseral bins. Misfit is data less model, in *MLT*/dipole-latitude. Note that the *X*- and *Z*-component data have been shifted in *MLT*, relative to the *Y*-component data, for clarity. *X*, *Y* and *Z* are spherical components and are dipole north, east and vertically down.

production of geomagnetic and solar activity indices and data. The OMNI data were obtained from the GSFC/SPDF OMNIWeb interface at http://omniweb.gsfc.nasa.gov and ACE level 2 data from http://www.sec.noaa.gov. Intermagnet observatory data were obtained from http://www.intermagnet.org. We would also like to thank our colleagues within the UK Universities GEOSPACE consortium for their help, comments and advice. This paper is published with the permission of the Executive Director, BGS (NERC). AWPT was supported by NERC grant NER/O/S/2003/00677. The referees are thanked for comments which significantly improved the original manuscript.

REFERENCES

Garrard, T.L., Davis, A.J., Hammond, J.S. & Sears, S.R., 1998. The ACE Science Center, *Space Sci. Rev.,* **86,** 1–4.

Iyemori, T., 1990. Storm-time magnetospheric currents inferred from mid-latitude geomagnetic field variations, *J. Geomag. Geoelec.,* **42,** 1249–1265.

King, J.H. & Papitashvili, N.E., 2004. Solar wind spatial scales in and comparisons of hourly Wind and ACE plasma and magnetic field data, *J. Geophys. Res.,* **110**(A2), A02209, doi:10.1029/2004JA010804.

Langel, R.A. & Estes, R.H., 1985. Large-scale, near-field magnetic fields from external sources and the corresponding induced internal field. *J. Geophys. Res.,* **90,** 2487–2494.

Langel, R.A. & Hinze, W.J., 1998. *The Magnetic Field of the Earth's Lithosphere–the Satellite Perspective*, Cambridge University Press, Cambridge.

Lesur, V., Macmillan, S. & Thomson, A., 2005a. A magnetic field model with daily variation of the magnetopsheric field and its induced counterpart in 2001, *Geophys. J. Int.,* **160,** 79–88.

Lesur, V., Macmillan, S. & Thomson, A., 2005b. The BGS magnetic field candidate models for the 10th generation IGRF, *Earth Planets Space,* **57,** 1157–1163.

Lowes, F.J., 1966. Mean square values on the sphere of spherical harmonic vector fields. *J. Geophys. Res.,* **71,** 2179–2180.

Maus, S. & Weidelt, P., 2004. Separating the magnetospheric disturbance magnetic field into external and transient internal contributions using a 1D conductivity model of the Earth, *Geophys. Res. Lett.,* **31,** L12614, doi:10.1029/2004GL020232.

Maus, S. & Lühr, H., 2005. Signature of the quiet-time magnetospheric magnetic field and its electromagnetic induction in the rotating Earth, *Geophys. J. Int.* **162,** 755–763, doi:10.1111/j.1365-246X.2005.02691.x

Maus, S., Rother, M., Hemant, K., Stolle, C., Lühr, H., Kuvshinov, A. & Olsen, N., 2006a. Earth's lithospheric magnetic field determined to spherical harmonic degree 90 from CHAMP satellite measurements, *Geophys. J. Int.,* **164,** 319–330, doi:10.111/j.1365-246X.2005.02833.x.

Maus, S., Rother, M., Stolle, C., Mai, W., Choi, S., Lühr, H., Cooke, D. & Roth, C., 2006b. Third generation of the Potsdam Magnetic Model of the Earth (POMME), *Geochem. Geophys. Geosyst.,* **7,** Q07008, doi:10.1029/2006GC001269.

Menvielle, M. & Berthelier, A., 1991. *Reviews of Geophysics,* **29,** 415–432.

Menvielle, M. & Berthelier, A., 1992, *Rev. Geophys.,* **30,** 91–92.

Olsen, N., 2002. A model of the geomagnetic field and its secular variation for epoch 2000 estimated from Ørsted data, *Geophys. J. Int.* **149,** 454–462.

Olsen, N., Sabaka, T.J. & Lowes, F., 2005. New parameterization of external and induced fields in geomagnetic field modelling, and a candidate model for IGRF 2005, *Earth Planets Space,* **57,** 1141–1149.

Olsen, N., Lühr, H., Sabaka, T.J., Mandea, M., Rother, M., Toffner-Clausen, L. & Choi, S., 2006. CHAOS—a model of the Earths magnetic field derived from CHAMP, Ørsted, and SAC-C magnetic satellite data, *Geophys. J. Int.,* **166**, 67–75, doi:10.1111/j.1365-246X.2006.02959.x.

Ritter, P., Lühr, H., Maus, S. & Viljanen, A., 2004. High-latitude ionospheric currents during very quiet times: their characteristics and predictability, *Ann. Geophys.* **22,** 2001–2014.

Sabaka, T.J., Olsen, N. & Purucker, M.E., 2004. Extending comprehensive models of the Earth's magnetic field with Ørsted and Champ data, *Geophys. J. Int.,* **159,** 521–547.

Stauning, P., Primdahl, F., Christiansen, F. & Watermann, J., 2003. Detection of fine-scale field aligned current structures from Orsted, in *Proceedings of OIST-4*, ISBN 87-7478-484-6, Danish Meteorological Institute, DK-2100 Copenhagen, Denmark.

Sugiura, M., 1964. Hourly values of the equatorial *Dst* for the IGY, *Ann. IGY,* **35,** 9–45.

Troshichev, O.A., Andrezen, V.G., Vennerstrom, S. & Friis-Christensen, E., 1988. Magnetic activity in the polar cap—a new index, *Planet. Space Sci.,* **36,** 1095–1102.

Viljanen, A. & Hakkinen, L., 1997. IMAGE magnetometer network, in *Satellite-Ground Based Coordination Sourcebook*, p. 111., eds Lockwood, M., Wild, M.N., Opgenoorth, H.J., ESA publications, Noordwijk, **SP-1198**.

Utada, H., Koyama, T., Shimizu, H. & Chave, A.D., 2003. A semi-global reference model for electrical conductivity in the mid-mantle beneath the north Pacific region, *Geophys. Res. Lett.,* **30,** 1194–1198, doi:10.1029/2002GL016092.